

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2002, 21:30:39 : Search time 1944.37 Seconds  
(without alignments)  
12446.212 Million cell updates/sec

Title: US-09-719-017A-2

Perfect score: 1793  
Sequence: 1 gaatccccgttgacatca.....caatactcaatgccccgcg 1793

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST: \*

1: em\_estbda:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inu:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrc:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	139	7.8	888	12	AQ329770 nbxb0045P
C 2	137.4	7.7	1073	12	AF114216
C 3	104.4	5.8	544	12	AQ989631 RfC00197
C 4	104	5.6	655	12	AQ990982 RfC01832
C 5	100.6	5.6	1101	12	AF029514 AF029514
C 6	99	5.5	622	12	AQ329769
C 7	83.6	4.7	580	12	AQ160684
C 8	82.2	4.6	409	9	AW711275
C 9	76.4	4.3	418	9	AW713376
C 10	74.2	4.1	417	9	AW710819
C 11	73	4.1	330	9	BE092318
C 12	72.2	4.0	694	9	BE092318
C 13	72.2	4.0	707	9	AM948129
C 14	72.2	4.0	708	9	AM948148
C 15	72.2	4.0	730	9	AM948130
C 16	72.2	4.0	738	9	AM948138
C 17	71.8	4.0	633	9	AM948152

18	69	3.8	643	9	AM948147
19	66.8	3.7	359	9	AM710710
20	53.4	3.0	348	9	AM708803
21	53.4	3.0	360	9	AM714224
22	52.8	2.9	434	9	AF408416
23	52	2.9	532	10	BG278308
24	51.6	2.9	918	12	AF114188
25	49.2	2.7	730	12	BH390360
26	48.2	2.7	477	10	BE600147
27	47.8	2.7	412	9	AI329382
28	47.6	2.7	507	10	BF484942
29	46.6	2.6	419	10	BF072824
30	46.4	2.6	925	12	CNS0091P
31	46	2.6	417	10	BF473461
32	45.4	2.5	259	10	BM373215
33	45.4	2.5	294	10	BM443910
34	45.4	2.5	295	10	BM098909
35	45.4	2.5	462	12	AQ016380
36	45.4	2.5	498	9	AV942445
37	45.4	2.5	577	9	AV945848
38	45.4	2.5	633	9	AV944572
39	45.4	2.5	719	12	AG010512
40	45.2	2.5	539	10	BE604694
41	44.8	2.5	688	10	BG912542
42	44.6	2.5	516	10	BE434941
43	44.6	2.5	517	10	BI269359
44	44.6	2.5	520	10	BI269270
45	44.6	2.5	564	10	BE460888

#### ALIGNMENTS

RESULT 1  
AQ329770/c 888 bp DNA linear GSS 08-JAN-1999  
LOCUS nbxb0045P09r CUGI Rice BAC Library Oryza sativa genomic clone  
DEFINITION nbxb0045P09r. DNA sequence.

ACCESSION AQ329770  
VERSION AQ329770.1 GI:4121620

KEYWORDS GSS.

SOURCE Oryza sativa.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 888)

Wing, R.A. and Dean, R.A.

A BAC End Sequencing Framework to Sequence the Rice Genome  
Unpublished (1998)

COMMENT

Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293

Email: twing@clemson.edu  
Seq primer: GGAACACGCTATGACCATG  
Class: BAC ends  
High quality sequence start: 13  
High quality sequence stop: 225.  
Location/Qualifiers

#### FEATURES

Source

1..888  
/organism="Oryza sativa"  
/strain="Japonica"  
/cultivar="Nipponbare"  
/db\_xref="taxon:4530"  
/clone="nbxb0045P09r"  
/clone\_lib="CUGI Rice BAC Library"  
/tissue\_type="Leaf"  
/lab\_host="E. coli DH10B"  
/note="Vector: pBelocAcl1; Site 1: HindIII; Site 2:  
HindIII; Rice is one of two most popular grains in the

world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT 184 a 233 c 175 g 278 t 18 others

ORIGIN

Query Match 7.8%; Score 139; DB 12; Length 888;  
Best Local Similarity 58.2%; Pred. No. 1.6e-28;  
Matches 211; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 269 gctgcggcgctacccctccacgctcgtggtgagcagcgccctgctcgaataacag 328  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 414 GCTAACGGGATATCATATTAGGATAGAGATCTGAATAATGGGGAGCAGTTAGA 355  
QY 329 tgcgcgtactatgccaactgcctcgtgagcagtgacagattcaacgattccca 388  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 354 CGATTCACAGTACGAAAGAGCAGANGCCGAGATGGCCGCGATCGCGGATTCGCCGA 295  
QY 389 ggcgcagcagacctgtgatttccatgcacatgggtatagagagcagcgcgcaag 448  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 294 GGGCGCGCTCGGCACAGCGCATGTGATCACTGGGCTATAGGACCGACCCCGCAG 225  
QY 449 ccttaccctgggccaatgctcgtacgacgaaggcgagatcgtgtgctgcgtcgca 508  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 234 CCTCTACATCGCTCACTACTGATCATGATCGTGACAGATCGTACAGCGCGAGCAAA 175  
QY 509 actcaaacactaacatggtgagcagccgtgtgtgtgtaagttatgccgaattatgat 568  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 174 GCTCAAGGCAACCCATGTGAGCGGACCATCTTGGTGAAGGCGAGCGACCATTCGC 115  
QY 569 tctgtcgcagcagcagcgtggtgcgtcgtggtcgtggtggtggtggtggtggtggt 628  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 114 CGGCTGTGAGAGGACCTTGGCGGCTTGGATCGCTTCTGCTGGGACATTTGAACC 55  
QY 629 cttagagcaagtacgctgtactccagacgaagacatcacattgcgcgtg 682  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 54 GCTCACCAATATATGCTGTAGCGCCAGACGACGATTCATCTCGCGCATG 1

RESULT 2  
AF114216/c 1073 bp DNA linear GSS 29-AUG-2000  
LOCUS AF114216 Salmonella typhimurium LT2, Lambda DASH II Salmonella  
DEFINITION typhimurium genomic clone 45-T7, DNA sequence.  
ACCESSION AF114216  
VERSION AF114216.1 GI:5281264  
KEYWORDS GSS.  
SOURCE Salmonella typhimurium.  
ORGANISM Salmonella typhimurium.  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Salmonella.  
REFERENCE 1 (bases 1 to 1073)  
AUTHORS Wong, R.M.Y. and McClelland, M.  
TITLE End Sequences of Salmonella typhimurium LT2 Lambda DASH II Clones,  
Li-Cor  
JOURNAL Unpublished (1999)  
COMMENT Contact: McClelland M  
Molecular Biology

Sidney Kimmel Cancer Center  
3099 Science Park Road, San Diego, CA 92121, USA  
Email: mclelland@lifsc1.sdu.edu  
Class: shotgun.  
Location/Qualifiers

FEATURES  
source 1..1073  
/organism="Salmonella typhimurium"  
/strain="LT2"  
/db\_xref="taxon:602"  
/clone="45-T7"  
/clone\_11b="Salmonella typhimurium LT2, Lambda DASH II"  
/note="Vector: Lambda DASH II; sequenced using Li-Cor  
sequencer."

BASE COUNT 226 a 292 c 277 g 277 t 1 others

ORIGIN

Query Match 7.7%; Score 137.4; DB 12; Length 1073;  
Best Local Similarity 85.5%; Pred. No. 5e-28;  
Matches 153; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1275 cctgcagcatgtaagcttggtccaccctgaccctgcgactcgaagtgaacgc 1334  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 848 CCGGCGGCACTAGCGCGGTGTCCTCAGACCCATCGCAATCAGAAAGTGAACGC 789  
QY 1335 cgttagcgcagatgtagtctggtggtcccatcgagagtaaggacgcagcatca 1394  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 788 CGTAGGCGCGATGTGATGTGGGGGTCTCCCATGACGAGATAGGAATCGCAGGATCA 729  
QY 1395 aataaaggaagcagtcagtcgaagaactggccttcgtttatcgtgtgttcg 1453  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 728 AATMAAGCAAAAGGCCCTGCGAAGACAGGCGCTTTGTTTGTCTGTTCTGCCGC 670

RESULT 3  
LOCUS AQ989631 544 bp DNA linear GSS 14-AUG-2000  
DEFINITION Photorhabdus luminescens strain W14 M13 library  
ACCESSION Photorhabdus luminescens genomic clone PLG00197, DNA sequence.  
VERSION AQ989631  
KEYWORDS AQ989631.1 GI:9648225  
SOURCE GSS.  
ORGANISM Photorhabdus luminescens.  
Photorhabdus luminescens.  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
Photorhabdus.  
REFERENCE 1 (bases 1 to 544)  
AUTHORS ffrrench-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T.,  
Daborn, P.V., Bowen, D., and Blattner, F.R.  
TITLE A genomic sample sequence of the entomopathogenic bacterium  
JOURNAL Photorhabdus luminescens W14: potential implications for virulence  
MEDLINE Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)  
COMMENT 20378633  
Contact: ffrrench-Constant RH  
Department of Biology and Biochemistry  
University of Bath  
South Building, Bath BA2 7AY, UK  
Tel: (44) 1225 826621  
Fax: (44) 1225 826779  
Email: bsrifc@bath.ac.uk  
This is one of 2,122 random reads from the M13 library. For  
annotation of identified clones (BLASTX, BLASTN and mapping to E.  
coli K12 genome) please see ffrrench-Constant et al. 2000, Nucleic  
Acids Res.  
Seq primer: M13 Forward  
Class: shotgun.  
Location/Qualifiers

FEATURES  
source 1..544  
/organism="Photorhabdus luminescens"  
/strain="W14"  
/db\_xref="taxon:29488"  
/clone="PLG00197"  
/clone\_11b="Photorhabdus luminescens strain W14 M13

Query match	5.88; Score 104; DB 12; Length 655;
-------------	-------------------------------------

OY	1387	agycatcaataaaacgaaagc	1409
Db	817	AGGCATCAATTAGACGATGGC	839

```

RESULT      6
LOCUS       AO329769                      622 bp    DNA        linear   GSS 08-JAN-1999
DEFINITION  nbxb0045P09f CUGI Rice BAC Library Oryza sativa genomic clone
ACCESSION   AO329769
VERSION     AO329769.1  GI:4121619
KEYWORDS
SOURCE      GSS.
ORGANISM    Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretidae; Oryzaceae; Oryza.
REFERENCE   1 (bases 1 to 622)
AUTHORS    Wing,R.A. and Dean,R.A.
TITLE       A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL     Unpublished (1998)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCATCATAGGG
            Class: BAC ends

FEATURES             location/qualifiers
     source           1..622
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="nbxb0045P09f"
                     /clone_1fb="CUGI Rice BAC Library"
                     /tissue_type="leaf"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pBelosAC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumanathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

BASE COUNT      112 a          200 c          190 g          119 t          1 others
ORIGIN
Query Match      5.5%; Score 99; DB 12; Length 622;
Best Local Similarity 54.9%; Pred. NO. 4e-17;
Matches 195; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

 Oy  693  tcgcgtacaggaacaaaggccatcgcgtcaacgagtgtaacataggcttcctcgcaa 752
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db   2  tgcgcttCAAGAAGCAAGCTTTCGCCGCTCGGTCCGAAGCTCAACAATAAGCCCTACGCCAG 61

 Oy  753  atcatactgatgaagccagtgtctaacatcgccgccagcagcatgttgtcacccagag 812
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db   62  gtcttACGGGGCGAGAAATgcttgcgtagatcgccgcccctcgccgggtgattctccgga 121

```

Qy	813	acccgagacatcgcgcggaagtagtgaacacaaacgcctccctcgtcgaagtgggccgcgcgc	872
Db	122	ATGATGACAGCTGATGATGTGGCGATACCGAGAGAGCGACGCCATTTTCTGCGCGGGGGCGCGC	181
Qy	873	agttccatgatttttgcgcgcgcgcgcgcgcgcacatctgctctccatctcgcacacagatgac	932
Db	182	TATGGAGTATCTATGTCGCCGATGGCGCGCGCGCGATGCGCAACCAATCGCACTGAGAGAG	241
Qy	933	gaaggcctgatactatgcgcgactcgaacatggaagaatttgccttcgcgcgaagcgatcaac	992
Db	242	GAGGCCCTCGTATGTGGCGGACATGATCTGTGCATGTATGCTCAACGCCAAGACCGCGGCA	301
Qy	993	gaaccttgggcgcactactccaacaccgcgaagccacccgctctgctgaactcgcgcgcgcgc	1047
Db	302	GACCGACCGGACATTTATGCGCGACCGACAGAGCTGCCACGCTGCTTTTCAACCCGG	356

RESULT	7
AOI0684	580 bp DNA Linear GSS 09-SEP-1998
LOCUS	
DEFINITION	mgxb00006L04r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0006L04r, DNA sequence.
ACCESSION	AOI0684
VERSION	AOI0684.1 GI:3557673
KEYWORDS	GSS.
SOURCE	Magnaporthe grisea.
ORGANISM	Magnaporthe grisea Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes; Incertae sedis; Magnaporthaceae; Magnaporthe. 1 (bases 1 to 580) Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury R., Phillips, K., Sasinowski, M., Ming, R.A. and Dean, R.A. A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome Unpublished (1998)
JOURNAL	Contact: Dean RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson University, Clemson, SC 29634 Tel: 864 656 5737 Fax: 864 656 4293 Email: rdean@clemson.edu Seq primer: GGAAACACTATGACATCG Class: BAC ends
TITLE	High quality sequence stop: 357. location/qualifiers
FEATURES	1..580 /organism="Magnaporthe grisea" /strain="70-15" /db_xref="taxon:148305" /clone="mgxb0006L04r" /clone_lib="CUGI Rice Blast BAC library" /tissue_type="Protoplasts" /lab_host="E. coli DH10B" /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25x genome coverage. High density colony filters are available upon request."
BASE COUNT	132 a 176 c 159 g 113 t
ORIGIN	
Query Match	4.7%; Score 83.6; DB 12; Length 580;
Best Local Similarity	54.2%; Pred. No. 1e-12;
Matches 192; Conservative	0; Mismatches 159; Indels 3; Gaps 1;

Query Match	4.68;	Score 82.2;	DB 9;	length 409;
Best Local Similarity	51.58;	Pred. No. 2.2e-12;		
Matches 189; Conservative	0;	Mismatches 178;	Indels 0;	Gaps 0;

```

/organism="Neurospora crassa"
/strain="Strain 30-7 (bd; A)"
/db_xref="taxon:5141"
/clone="g8c04ne"
/clone_11b="Neurospora crassa evening cdna library"
/tissue_type="tissue harvested following 22hr growth in
dark"
/notes="Vector: pBluescript SK-; site_1: XbaI; site_2:
EcoRI; See: Bell-Perdersen D., et al. PNAS 93:13096, 1996
5' end of cdna cloned into XbaI site of pBluescript; 3'
end of cdna cloned into EcoRI site of pBluescript"
BASE COUNT      94      136 c      103 g      85 t

```

BASE COUNT	ORIGIN
94 a	136 c
103 g	85 t

Query Match	4.38;	Score 76.4;	DB 9;	Length 418;
Best Local Similarity	51.58;	Pred. No. 1e-10;		
Matches 176;	Conservative	0;	Mismatches 166;	Indels 0;
			Gaps	0;

OY	166	ggagagccgcgccttaacagagccgcgcctctcccaacacagatctcgagcaacaggggtgtgtataaanaac	205
Db	77	gcccacagcgggtactttccgaacccctccctcggttcgattctcgaaggagggggttcgcgaagac	136
OY	206	catgagctgctgcgtcagagccgcgcgatbagggctgtgactcgatcgtgtttgttgaanaac	265
Db	137	gattcacttccatttaacgaaagccgggtcagcccgccgctgcaagctcgttcgcctttcccaact	196
OY	266	cggagctgcgcgcgcgtatccctctcccaagctcgtgtgtagcgagccgagccctgtctgcataata	325
Db	197	ctggattcccccgccttattccattacttgatcttggaaggtcaccttaccagacagtccgttcgctt	256
OY	326	cagtgccgcgtactatgccaactgcgtctcgtctgtgacagtgcaagatttcaacgcataatgc	385
Db	257	gcttgaagaagatgcccgcggaacgccatggccgcttgactcgtgcagcagattccgttcgcattcag	316
OY	386	ccagagccgacagagacttcggtatcttctcgcacatgggtttatagcgagcgcaagcgcgag	445
Db	317	aaggcctgcctagagcaaacacacattttattcttcgctcggcttcgccgaataatgacaccaccc	376
OY	446	cagccttaccttgggcacatgcccgtatcagagcaagaagcca	487
Db	377	cacccttaccctggcgcaagctcttatcgtatccctccactgcgca	418

RESULT 10  
AM710819

**DEFINITION** *ebh01ne.f2* Neurospora crassa evenng cdna library Neurospora crassa cdna clone *ebh01ne 5'*, mRNA sequence.

KEYWORDS EST.  
SOURCE Neurospora crassa.

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales; Sordariaceae; Neurospora.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

1 (uses 1 to 411)  
 Zhu, H., Lai, H., Kupfer, D., Dunlap, J. C. and Roe, B. A  
 Two Neurospora crassa EST databases  
 Unpublished (1998)  
 Other ESTs: e6h0line.r1

contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
Department of Chemistry and Biochemistry  
Advanced Center for Genome Technology, University of Oklahoma  
620 Parrington Oval, Norman, OK 73019, USA

Email: broeeu@uconn.edu  
We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center  
Seq Primer: Universal Forward Primer  
High quality sequence, stop: 393.

FEATURES	Location/Qualifiers
source	1. .417

```
/organism="Neurospora crassa"  
/strain="Strain 30-7 (bd; A)"
```

```

/ad_xlet= lax011:
/clone="e6h01ne"

```

```
clone_lib="Neur
/tissue_type="ti
```

dark"

ECORI; See: Bell-Perdersen, D., et al. PNAS 93:13096

3' end of cpDNA cloned into EcoRI site of pBluescript";

a	136	c	100	g	85	t	1	others
---	-----	---	-----	---	----	---	---	--------

Query Match	4.1%	Score 74.2	DB 9	Length 417
Best Local Similarity	51.3%	Pred. No. 4.4e-10		
Matches 172; Conservative	0	Mismatches 163	Indels 0	Gaps 0

OY	146	ggagacgcgcgttaacagagcgcgcctctcccaactacagatctcgtgaaacgggtgtgtgataa	205
Db	82	gcccacagcgttacttctccgaacccctcctcgttccatctcgaaggagcgcgttcgcacac	141
OY	206	cattagactcgtgcgtcgaagccgcgcattgaaggctcgtgaactcgtatcogtltgtgtgaac	265
Db	142	gattcacatttcatttaacgaagcgggctcagccgcgcgtgcgaagctcgtcgcctttccgaagt	201
OY	266	ctggctgcgcgcgtatccctctccacgctcgtgcgttgaggcgaacgcgcgtcgtgtcgtgaata	325
Db	202	ctggattccccggcgtttccatctggaatgtggaaagtcaccttcacagactccctgccttat	261
OY	326	cagtcgcgcgtactatgcgaactcgtcctcgtgacagltcagaglttcaaagcattcgc	385
Db	262	gctgaagaagatraccccggaamcgcctatggcccttgactcgcgcgaagttccgtgcacatcag	321
OY	386	ccagacgcgaacgagacctcggatcttcaatcgcactcgtggttataagagcgcgaacgcgcgc	445
Db	322	aaggcctcgtcgaaggaacacacagattttatgtctcgcctcggccttcggcgaaatgcacaccc	381
OY	446	cagccttactcgtggccaatgcctcgtatcgaagaca	480
Db	382	caccctttacctggcgcaagctcgtatcagatccca	416

RESULT 11  
BE092318/c

ACCESSION	BE092318	
VERSION	BE092318.1	GI:8482770

SOURCE	human.
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**AUTHORS**

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldmann, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Rucher, P., Tonangeel, C. V., O'Hare

M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
-------	--

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research  
Laboratory of Cancer Genetics  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo-SP.

Brazil  
Tel: +55-11-2704922

Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry may be used in the following manner:

project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6tt2=IL2-BT0734-240>  
A00-072-D10f3=2000-04-24c4=1)

400 0/2 DDBS3-2000 04 29864-1)  
Seq primer: puc 18 forward  
High quality sequence stop: 292.

FEATURES	Location/Qualifiers
source	1.330

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"
```

```
/clone_11b="B10/34"  
/dev stage="Adult"
```

```

source
1. .694
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="MT0013"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: Sma1; Site_2:
Sma2; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent Application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low

```

Smal: A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT	168 a	166 c	200 g	173 t
ORIGIN				
Query Match	4.0%; Score 72.2; DB 9; Length 707;			

DB	Accession	Similarity	Score	DB	Length	Gaps
Best Local	Similarity	78.9%	Pred. No. 2.1e-09;	Indels	0;	Gaps 0;
Matches	86;	Conservative	0;	Mismatches	23;	Indels 0;
OY	1289	agcttgagtcaccactgaccccatgcccgaatccagaagtgaacgccgtatgcccgaatg	1348			
Db	563	AGCATTGGACCACTGTATCCATCCCGAAGTACACAGCAAGAAACATGATGCATGCCGATG	504			
OY	1349	taagtgtggagtcctcccatgcccgaagtgaaggaacgcgcgcgcataaat	1397			
Db	503	TAGTGTGGCGTTTCCCATGTCAGAGTACGTATGCTCAAGATTAAATT	455			
RESULT	14					
LOCUS	AM948148/c	708 bp	mRNA	linear	EST 31-MAY-2000	
DEFINITION	RCO-MT0013-280300-031-e04 MT0013 Homo sapiens cDNA, mRNA sequence.					
ACCESSION	AM948148					
KEYWORDS	AM948148.1 GI:8125922					
VERSION	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
AUTHORS	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.					
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)					
MEDLINE	20202653					
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-3704922 Fax: +55-11-3707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=et2-RCO-MT0013-280300-031-e04&t3=2000-03-28&t4=1) Seq primer: puc 18 forward High quality sequence start: 65 High quality sequence stop: 676.					
FEATURES	location/Qualifiers					
source	1..708					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone_lib="MT0013"					
	/dev_stage="Adult"					
	/note="Organ: marrow; Vector: puc18; Site:1: Sma1; Site:2: Sma1; A mini-library was made by cloning products derived from ORESSES PCR (U.S. Letters Patent application No. 196 ,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."					
BASE COUNT	164 a 172 c 201 g 171 t					
ORIGIN						
Query Match	4.0%; Score 72.2; DB 9; Length 708;					
Best Local Similarity	78.9%; Pred. No. 2.1e-09;					
Matches	86; Conservative 0; Mismatches 23; Indels 0; Gaps 0;					
OY	1289	agcttgagtcaccactgaccccatgcccgaatccagaagtgaacgccgtatgcccgaatg	1348			
Db	561	AGCATTGGACCACTGTATCCATCCCGAAGTACACAGCAAGAAACATGATGCATGCCGATG	502			

Oy	1349	tagttggaggtctcccatgccccatgagagtagggagacgcgcgcatacaat	1397
Db	501	TAGTGTGGGGTTTCCCATGTGAGAGTAGTCATGCTCAAGATTAAATT	453
RESULT	15		
LOCUS	AM948130	730 bp	mRNA
DEFINITION	RCO-MT0013-280300-031-D09 MT0013 Homo sapiens cDNA, mRNA sequence.		EST 31-MAY-2000
ACCESSION	AM948130		
VERSION	AM948130.1	GI:8125904	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 730)		
AUTHORS	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. J., Zago, M. A., Bordin, S., Costa, F. F., Goldstein, G. H., Carvalho, A. F., Matsukuma, A., Bala, G. S., Simpson, D. H., Brunstein, A., deoliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
MEDLINE	200202663		
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?c1=4t2-RCO-MT0013-280300-031-b09&t3=2000-03-28&t4=1) Seq primer: puc 18 forward High quality sequence start: 9 High quality sequence stop: 666. Location/Qualifiers 1..730 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="MT0013" /dev_stage="Adult" /note="Organ: marrow; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESRES PCR (U.S. Letters Patent application No. 199716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
FEATURES	source		
	BASE COUNT	171 a	171 c
	ORIGIN	206 g	182 t
	Query Match	4.0%;	Score 72.2;
	Best Local Similarity	78.9%;	DB 9;
	Matches	86;	Conservative 0;
		Mismatches 23;	Indels 0;
		Gaps 0;	
Oy	1289	agcttgagttccacactgccccatgccccatgacacagaagtgaaagccgtagagccgaatg	1348
Db	568	AGCATTGGAACCACTGTGATCCATCCGCAACACACAGTAACATGCATGCGCGATGG	509
Oy	1349	tagttggaggtctcccatgccccatgagagtagggagacgcgcgcatacaat	1397
Db	508	TAGTGTGGGGTTTCCCATGTGAGAGTAGTCATGCTCAAGATTAAATT	460



Tue Sep 24 19:11:21 2002

us-09-719-017a-2.rst

Page 9

Search completed: September 17, 2002, 00:35:53  
Job time: 11114 sec

---

